Remarks

Claims 1-31 are pending in the application. Claims 1-31 are rejected. All rejections are respectfully traversed.

Claim 23 is rejected under 35 U.S.C. 112, second paragraph, as being indefinite.

Claim 22 is amended to define the coordinates $o(u, v)$ for the pixels in the output image used in claim 23, see paragraph [095]. Claim 1, line 10, provides antecedent basis for "viewing parameters" in claim 23.

8. Claims 1-4, 14, 16, 20-24, and 27-31 are rejected under 35 U.S.C. 103(a) as being unpatentable over Anderson (US 5, 714, 997) in view of Wenger et al (US 2004/0179591).

As stated in the present application at paragraph [058], the system provides stereoscopic color images for multiple viewpoints without special viewing glasses. The claimed invention includes a three-dimensional display unit configured to concurrently display a plurality of output videos onto a single display surface according to viewing parameters. In contrast, the Anderson system uses a pair of stereo goggles having separate right-eye and left-eye displays. Anderson cannot use a single display surface.

As per claim 2, the system uses a plurality of cameras to acquire calibration images displayed on the display surface of the three-dimensional display

unit to determine the viewing parameters. The cameras 112 described by Anderson with respect to Figure 7 acquire videos of the scene or coverage area 40, and not the display surface as claimed. Anderson describes calibrating those cameras. The cameras as claimed acquire images of calibrations images that are used to determine viewing parameters.

As per claim 3, Anderson uses a video output device in the form of a pair of stereo goggles having separate right-eye and left-eye displays, see Figure 31 and the description thereof. The system of Wenger projects non-overlapping side-by-side images of sub-scenes, see Figures 1-4, 6-9 and 11. Such a system would not work with the Anderson goggles. The Anderson stereoscopic goggles require that the left and right eye see approximately the same view to obtain the virtual reality effect, see the description of Figure 32 at column 39:

> FIG. 32 illustrates, as an example, certain parameters used by the right and left renderers, 2302 and 2303, in rendering their respective right-eye and left-eye frames. For each eye, the viewpoint is respectively offset from the current virtual viewpoint by half the interocular distance (i.e., distance between the viewer's eyes), and directed towards an object being focused upon. For example, the right eye viewpoint 2402 is offset from the virtual viewpoint 2401 by a distance j_{RIGHT} and directed towards a point of convergence 2407 determined by an object being focused upon, and the left eye viewpoint 2403 is offset from the virtual viewpoint by a distance j_{LEFT} and directed towards the point of convergence 2407 determined by the object being focused upon. The corresponding viewing directions for the right and left eye frames are preferably calculated by the right and left renderers 2302 and 2303 to converge on, for example, an image surface of closest depth which is not part of the background.

If the Wegner images were combined with the Anderson goggles, each eye would see a different sub-scene, and at most two sub-scenes can be used without the 3D effect. With all due respect, the Examiner's statement,

> The advantage is that it adjusts the picture quality for each of a plurality of images that comprise a picture and that distributes or balances the transmission of the plurality of images in order to optimize the overall picture quality in a video transmission.

is a *non sequitur*. Projecting different images to different eyes using the Wegner projectors and the Anderson goggles has nothing to do with distributing and balancing the quality of video transmission. The combination of Wegner and Anderson would be a virtual mess and not a virtual reality.

Organic LEDs as claimed are incompatible with the stereo goggles of Anderson.

With respect to claim 14, as stated above the systems of Anderson and Wegner are incompatible with each other.

With respect to claim 21, each video camera has viewing parameters that include a position, orientation, field-of-view, and focal plane. The Examiner states that Anderson does not teach parameters, but that it is "well known in the art" to use these viewing parameters. The Examiner provides no support for this assertion. MPEP 2144.03 states:

> "It would <u>not</u> be appropriate for the examiner to take official notice of facts without citing a prior art reference where the facts asserted to be well known are not capable of instant and unquestionable demonstration as being well-known. For example, assertions of technical facts in the areas of esoteric technology or specific knowledge of the prior art must always be supported by citation to

> some reference work recognized as standard in the pertinent art. …; *In re Eynde*, 480 F.2d 1364, 1370, 178 USPQ 470, 474 (CCPA 1973) ('[W]e reject the notion that judicial or administrative notice may be taken of the state of the art. The facts constituting the state of the art are normally subject to the possibility of rational disagreement among reasonable men and are not amenable to the taking of such notice.')."

3D television is still considered an esoteric technology. 3D television is still very much in the experimental stage, and commercial 3D television systems are not yet available. The field of 3D television is still being explored and what type of viewing parameters are required to provide plausible 3D images to viewers is very much an open question. Any number of factors could be considered. Certainly, the viewing parameters as claimed are not capable of such a demonstration. Certainly, that the viewing parameters should include position, orientation, field-of-view, and focal plane of each camera to accurately provide a 3D viewing experience is not *instantly and unquestionably* demonstrable. 3D television acquisition and rendering is a technically difficult area and assertions that certain procedures in the field are well known must be supported by evidence as required by the MPEP.

Furthermore, MPEP 2144.33 goes on to state:

> "If such notice is taken, the basis for such reasoning must be set forth explicitly. The examiner must provide specific factual findings predicated on sound technical and scientific reasoning to support his or her conclusion of common knowledge. See *Soli*, 317 F.2d at 946, 37 USPQ at 801; *Chevenard*, 139 F.2d at 713, 60 USPQ at 241. The

applicant should be presented with the explicit basis on which the examiner regards the matter as subject to official notice and be allowed to challenge the assertion in the next reply after the Office action in which the common knowledge statement was made."

The Examiner sets forth no reasoning as to why he believes that viewing parameters as claimed are well known. The Examiner's statements:

However, it is well known in the art to use viewing parameters that include a position, orientation, field-of-view, and focal plane for a video camera (Official Notice).

and

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention was made to modify the virtual reality television of Anderson with the method of Wenger to apply viewing parameters such as position, orientation, field of view and focal plain of video cameras.

are merely conclusions with no reasoning as to why the Examiner feels that these are well known. As such, the Examiner is depriving the Applicants of the ability to challenge his assertion, which is clearly in violation of the directives of the MPEP.

The Examiner is respectfully requested to either provide evidentiary support and to clearly state his rationale in support of his assertions or to withdraw his rejections.

As per claim 22, the controller determines, for each output pixel $o(u, v)$ in the output video, a view number v and a position of each source pixel $s(v, x,$

y) in the decompressed videos that contributes to the output pixel in the output video.

The Examiner refers to paragraphs [040], [041] and [043] of Wenger.

Paragraph [040] describes sending data associated with images to the encoders. The method for the transmission of the video images can be by the means of either an analog or digital connection. A Bit Rate Control Manager receives a QP value for each encoder, and generates a new bit rate for each encoder. The bit streams (with bit rates as determined by the BCM) are then multiplexed with a multiplexer and distributed over a wide area network to a remote station.

Paragraph [0041] describes a remote station 36. Four bit streams are received by a plurality of decoders. The typical video conferencing environment also permits images to be captured at the remote station for transmission back to the first station.

Paragraph [0043] describes how each of the cameras is coupled to each digital signal processor and synchronized. All digital signal processors communicate to each other the target bit rate that was used when coding the last picture captured by the cameras. The DSPs also communicate to each other the achieved quality in the form of the average QP value for that last picture captured by the cameras. Each digital signal processor runs an identical algorithm in order to calculate the next target bit rate. This target or new bit rate is used in the encoding step of the next picture captured by its

respective camera, as the rate control parameter. Using the newly determined bit rate, the system and method cause the average QP value to be substantially the same, thereby insuring a substantially identical quality of the pictures that make up the broadband image. This process continues during the lifetime of a video conference so that the quality of the broadband image displayed at the station is of the best possible quality that the bandwidth of the network and various components of the system will allow. Note that the average QP is calculated by adding all QP values of all macroblocks belonging to the coded picture, and dividing it by the number of macroblocks in the coded picture. Note also, that the calculation of the new target bit rate can also be performed for a plurality of macroblocks while coding a picture, an entity called a slice in many video compression standards. Doing so, the accuracy of the method and system can be higher, but the computational demands are lower. It is also possible to increase the time interval between the calculations of the new target bit rate to a multitude of the inverse of the frame rate. In this case, the computational demands are lower. The quality measurement such as in the form of the average QP, is always performed using all those macroblocks that were coded in the time interval between the last calculation and the current calculation.

None of the above three paragraphs indicate how source pixels are mapped to output pixels. The rejection of claim 22 is improper.

As per claim 23, the Examiner merely relies on the improper rejection of claim 22, and does not consider the additional limitations of claim 23, which

considers a weighted linear combination of *k* source pixels according to

$$o(u,v) = \sum_{i=0}^{k} w_i s(v,x,y)$$

, where blending weights w_i are predetermined by the controller based on the viewing parameters. Paragraphs [0040-0043] do not describe a weighted linear combination of pixels. The rejection of claim 23 is improper since the Examiner fails to provide any prior art that teaches the claimed limitations.

At paragraph [0038] Wenger states:

> In one embodiment of this invention described later herein, the average quantization factor or QP value of all macroblocks is also used heuristically as an indication of the quality of one of the coded pictures that are optimized by the MCRC algorithm. Note that the average QP is calculated by adding all QP values of all macroblocks belonging to the coded picture, and dividing it by the number of macroblocks in the coded picture.

There is nothing there about a block of source pixels contributing to each output pixel.

As per claim 27, claimed is an arrangement of the cameras and an arrangement of the display units, which are substantially identical, and the number of cameras and display units is greater than two. Anderson is restricted to a single display unit with left and right eye display generator, see Figure 31. Such a system is incapable of operating with more than two display units.

Wenger cannot be combined with Anderson. Anderson uses the principles of stereo television systems that use left/right camera pairs to generate separate

images to be displayed to the left and right eyes of a viewer to provide an illusion of depth; viewing position is still restricted to actual camera locations. Wenger uses a compositing of non-overlapping images, see Figure 1b and paragraph [0013]: "The projector beam directions of all projectors P1 to P4 are arranged in such a way that the four displayed sub-images I1 to I4 spatially compose a full image I that geometrically resembles the captured scene A."

As per claim 29, Wenger acquires a wide-band scene. As best as can be understood, in Wenger a video of a "wide-band scene A, with an aspect ratio of 16:3, consists of four spatially adjacent sub-scenes A1, A2, A3, and A4." That is, the "wide-band" applies to the *spatial* arrangement where the width is much greater than the height. The invention acquires high dynamic light-fields. A dynamic light field varies rapidly in the *temporal* domain. The Wenger spatial arrangement has nothing to do with the claimed temporal variation.

Similarly, Wenger deals with the spatial arrangement of the projected images, see paragraph [0013]:

> The projector beam directions of all projectors P1 to P4 are arranged in such a way that the four displayed sub-images I1 to I4 spatially compose a full image I that geometrically resembles the captured scene A.

The dynamics of the claimed rendering has to do with temporal variation.

9. Claims 5-13 are rejected under 35 U.S.C. 103(a) as being unpatentable over Anderson (US 5, 714, 997) in view of Wenger et al (US 2004/0179591), and in further view of Tung (US 2002/0122145).

As per claims 8-9, Tung states:

> [0039] While the preferred embodiment has been described using linearly polarized material with each portion of the system the use of circularly polarized elements is equally useful. In such a system, the input linear polarizer **32** in the preferred embodiment, would be replaced by a passive circular polarizer. The LC shutter glass assemblies in the preferred embodiment would have, in addition to the active rotator **20** and the linear polarizer **22** (nearest the eye), a quarter wave element in the optical path at the location where the second polarizer existed in **FIG. 1**.

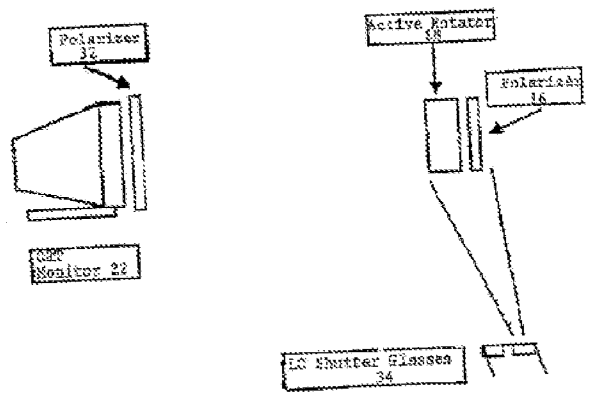
Tung does not describe a passive display device that uses a flexible fabric as claimed.

As per claim 10, the viewing direction of the viewer is not described by Tung:

> [0036] The use of liquid crystal shutter glasses for 3D stereoscopic viewing is illustrated in **FIG. 560**. A sequence of images that alternate between the left and right view perspective is displayed on a viewing screen. The screen is typically cathode ray tube (CRT)-based display, or a CRT-based projector. However, as will be discussed below, the display device may also be either a direct projection or rear projection device. Two shutters **62** and **64**, serving as the primary optical components of the shutter glasses, are opened and closed so that the left eye shutter is open when the left eye image **66** is displayed and the right eye shutter is open when the right eye image **68** is displayed. When a shutter is closed, ideally all light is blocked from passing through the shutter element as shown in the figure. When the shutter is opened, the shutter is transparent allowing the underlying eye to see the intended image. **FIG. 5** illustrates the transition from the left eye view to the right eye view from left to right with the left eye cycle on the left of the figure and the right eye cycle on the right of the figure. In the figure, time increases from left to right.

As per claims 10-12, paragraph [0036] above does not describe a display surface of a display unit disappearing when images are displayed on the surface.

As per claim 13, Figure 2 of Tung and paragraph [0036], see above, of Tung do not describe acquiring images. The Figure only shows a CRT and lenses.



10. Claims 15 and 17-19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Anderson (US 5, 714, 997) in view of Wenger et al (US 2004/0179591), and in further view of Haskell et al (US 6,055,012).

As per claim 15, the claimed system has a plurality of video cameras that are in a regularly spaced linear and horizontal array. Haskell shows a bank of cameras each capturing a slightly different view of the scene. The cameras are assumed to be located side-by-side, however, in all generality, the cameras may be located in any configuration and each camera may have any orientation. Figure 2 shows cameras in a side-by-side configuration. The cameras are shown as a one-dimensional array or they can be in a two-

dimensional array. Haskell does not describe cameras that are regularly spaced.

As per claim 17, in the claimed system, an optical axis of each video camera is perpendicular to a common plane, and the up vectors of the plurality of video cameras are vertically aligned. Haskell does not describe cameras with perpendicular optical axes. Instead, Haskell describes parallel optical axes:

> Moreover, there is no implied constraint to a parallel optical axis on the camera orientation geometry. While the arrangement of views can be application dependent, and thus very flexible, efficient coding can only be performed if there is significant correlation between the views.

As per claim 18, the viewing parameters include intrinsic and extrinsic parameters of the video cameras. There is no description of extrinsic parameters at column 6.

> DETAILED DESCRIPTION
>
> FIG. 1 shows a multi-view video imaging, coding/decoding and display system according to our invention to illustrate in detail, the framework and method of this invention which yields efficient compression of multi-view video while ensuring complexity constrained encoding/decoding and compatibility with normal video displays. A scene **100** is captured by a bank of cameras (or other imaging sensors) **105**, where camera **101**, **102**, **103** and **104** capture corresponding different views of the scene and output them on lines **111**, **112**, **113** and **114**. For the purposes of discussion we have assumed four cameras, each capturing a slightly different view of the scene, although in general, a camera may have multiple lens and be able to capture more than one view. Also, the cameras are assumed to be located side-by-side, however, in all generality, the cameras may be located in any configuration and each camera may have any orientation.

Similarly, there is no description of extrinsic parameters at column 7. As

known in the art, extrinsic parameters of a camera describe the position and orientation of the camera using a rotation matrix **R** and a translation matrix **t**. There are three parameters for the rotation, another three for the translation, see paragraph [078]. The claimed viewing parameters extend extrinsic parameters to 3D television system. Haskell does not teach, show or suggest extrinsic parameters for 3D television, at columns 6 and 7, or anywhere else.

As per claim 19, the claimed system includes means for selecting a subset of the plurality of cameras for acquiring a subset of videos. The Examiner cites Wegner col 7, line 40, col 8 lin 5:

> **[0040]** A plurality of cameras C1, C2, C3 and C4 capture individual images **20**, **22**, **24** and **26** and sends the data associated with the images **20-26** to the encoders (**En1-En4**) **12-18**, respectively. The method for the transmission of the video images can either be by the means of an analog or digital connection. The DSPs **12*a*-8*a*** each comprise Bit Rate Control Manager process (BCM), schematically illustrated as block **30** in **FIG. 2**, which receives a QP value for each encoder **En1-En4** and actual bit rate for each image captured by cameras C1-C4, respectively, and generates a new bit rate for each encoder in the manner described later herein. The bit streams (with bit rates as determined by the BCM **30**) are then multiplexed with a multiplexer **32** and distributed over a wide area network **34** to a remote station **36**.
>
> **[0041]** **FIG. 3** illustrates the receiving or remote station **36**. The multiplex signal is received over network **34** and received by demultiplexer **38** which demultiplexes the signal into four bit streams in the illustration. The four bit streams are received by a plurality of decoders **DEC140**, **DEC242**, **DEC344** and **DEC446** each having an associated digital signal processor **40*a***, **42*a***, **44*a*** and **46*a***, respectively. For ease of illustration and understanding, the invention is described using encoders **En1-En412-18** and decoders **DEC1-DEC440-46**, but it should be understood that a fewer or greater number can be used, provided there is more than just one. Also, the invention can be used in a combined integrated unit encoder/decoder commonly known as "codecs." For example, it should be understood that while the remote station is capable of receiving captured images from the first station **11** illustrated in **FIG. 2**, the typical video conferencing environment will also permit images to be captured at the remote station **36** for transmission back to the first station **11**.
>
> **[0042]** It should be understood that the bit rate corresponds to the bit rate corresponding to the images captured at the first station **11**. Moreover, note that the QP value for each of the captured images **48**, **50**, **52** and **54** is identical because it is a part of the bit stream itself. In the illustration being shown in **FIGS. 2 and 3**, the QP value of the images **48-54** displayed is illustrated as being **10**, and the bit rates are 5 kbit/s for En1, 10 kbit/s for Enc2, 10 kbits for En3, and 0 kbit/s for En4.

[0043] Referring to FIG. 2, it should be understood that each of the cameras C1, C2, C3 and C4 are coupled to each digital signal processor (DSP) 12*a*, 14*a*, 16*a*, and 18*a*. It should be further understood, however, that the cameras C1-C4 are synchronized in time by externally providing a horizontal and vertical synchronization signal and pixel

There is nothing about selecting cameras in the above paragraphs.

11. Claims 25 and 26 rejected under 35 U.S.C. 103(a) as being unpatentable over Anderson (US 5, 714, 997) in view of Wenger et al (US 2004/0179591), and in further view of Morishita (US 4, 737, 840).

Claims 25 and 26 specify a three-dimensional display unit that includes a display-side lenticular sheet, a viewer-side lenticular sheet, and a diffuser. A substrate can be between each lenticular sheets and the diffuser or between the lenticular sheets and the reflector.

First as stated above, it should be noted that Anderson uses a video output device in the form of a pair of stereo goggles having separate right-eye and left-eye displays, see Figure 31 and the description thereof. The Examiner states:

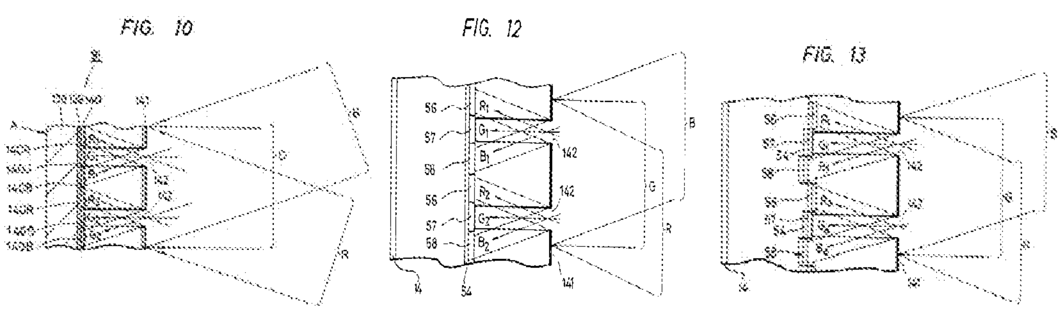
> Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention was made to modify modified the virtual reality television of Anderson according to Morishita's invention to incorporate features recited in claim 26 because it provides brighter pictures and widens the range of viewing angles.

It should be obvious to one of ordinary skill in the art that the system as described by Anderson can never work with a lenticular lens as described by Morishita. The combination of Morishita and Anderson results in an

inoperable system. The separate view goggles of Anderson cannot incorporate the lenticular lens of Morishita.

Furthermore, even if Morishita could be combined with Anderson, Morishita only describes a single lenticular lens. Claimed are two lenticular sheets, one is a display-side lenticular sheet, and the other a viewer-side lenticular sheet. There is nothing like that in Morishita.

As shown in Figures 10-13, the substrate 141 is neither between the lenticular sheets and the diffuser nor between the lenticular sheets and the reflector. The substrate 141 in Morishita in all embodiments is on an outside surface.



In addition, Morishita does not show or describe a reflector.

It is believed that this application is now in condition for allowance. A notice to this effect is respectfully requested. Should further questions arise concerning this application, the Examiner is invited to call Applicants' attorney at the number listed below. Please charge any shortage in fees due in connection with the filing of this paper to Deposit Account <u>50-0749</u>.

Respectfully submitted,
Mitsubishi Electric Research Laboratories, Inc.

By

/Dirk Brinkman/

Dirk Brinkman
Attorney for the Assignee
Reg. No. 35,460

201 Broadway, 8th Floor
Cambridge, MA 02139
Telephone: (617) 621-7517
Customer No. 022199